IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT:	Adriana Dumitras and Barin G. Haskell
APPLICATION NO.:	10/655,564
FILING DATE:	September 3, 2003
TITLE:	Look-Ahead System and Method for Pan and Zoom Detection in Video Sequences
EXAMINER:	Christopher G. Findley
GROUP ART UNIT:	2621
ATTY. DKT. NO.:	18602-08204

CERTIFICATE OF ELECTRONIC (EFS-WEB) TRANSMISSION

I hereby certify that this correspondence is being transmitted via the Office electronic filing system in accordance with 37 C.F.R. § 1.8(a)(i)(C) from the Pacific Time Zone of the United States on the local date shown below.

Dated: December 3, 2008 By: /Brian G. Brannon/
Brian G. Brannon, Reg. No. 57,219

COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

APPEAL BRIEF

Real Party in Interest

The subject application is owned by Apple Inc. of Cupertino, California.

Related Appeals and Interferences

There are no known related appeals or interferences that may directly affect, be directly affected by, or have a bearing on the Board's decision in the pending appeal.

Status of Claims

Claims 1-18 stand finally rejected. On August 4, 2008, the final rejection of claims 1-18 was appealed. The claims on appeal are set forth in an appendix attached hereto.

Status of Amendments

Appellant has not amended the claims since the final rejection.

Summary of Claimed Subject Matter

The independent claims on appeal are claims 1, 7 and 13. Claim 1 recites a method for detecting at least one of a pan and a zoom in a video sequence (e.g., FIG. 3; spec. page. 9, ¶ [0030]). A set of frames are selected from a video sequence (e.g., FIG. 3, steps 302, 304; spec., page 9, ¶ [0030]). For each frame in the selected set of frames, a set of motion vectors are determined and a motion angle for each motion vector is also determined (e.g., FIG. 3 steps 306, 308; spec., pages 9-10, ¶ [0031]). In each frame, at least two largest regions having substantially similar motion angles are identified (e.g., FIG. 3, step 310; spec., page 10, ¶ [0033]). The percentages of each frame covered by the at least two largest regions and a statistical measure of the motion angles for at least one of the two largest regions are determined (e.g., FIG. 3, steps 312, 314; spec., page 11, ¶ [0035]). The percentages of each frame covered by the at least two largest regions and the statustical measure of the at least two largest regions are compared to threshold values to identify a pan or a zoom in the video sequence (e.g., FIG. 3, steps 318, 320; spec., pages 11-12, ¶ [0036]).

Claim 7 recites a system for detecting at least one of a pan and a zoom in a video sequence (e.g., FIG. 2, video encoder 200; spec., page 8, ¶ [0028]). The system includes a prepocessor for selecting a set of frames from a video sequence (e.g., FIG. 2, preprocessor 202; spec., page 9, ¶ [0029]). Also included in the system is a motion analyzer which determines a set of motion vectors for each frame in the selected set of frames and determines a motion angle for each motion vector (e.g., FIG. 2, motion analysis block 208; spec., page 9, ¶ [0029]. The motion analyzer also identifies at least two largest regions in each frame having motion vectors with substantially similar motion angles and determines the percentages of each frame covered by the at least two largest regions (e.g., FIG. 3 steps 306, 308, 310; spec., pages 9-10, ¶ [0031], [0033]). A statistical measure of the motion angles for at least one of the two largest regions is determined by the motion analyzer, which also identifies a pan or a zoom in the video sequence by comparing the percentages of each frame covered by by the at leat two largest regions and the determined statistical measure to threshold values (e.g., FIG. 3, steps 314, 318, 320; spec., pages 11-12, ¶ [0035], [0036]).

Claim 13 recites a computer-readable medium which stores instructions that when executed by a processor in a system for detecting a pan or a zoom cause the processor to perform a set of operations for detection a pan or a zoom (e.g., FIG. 3; spec. pages 7-8, ¶ [0025]; page 9, ¶ [0030]). The operations cause selection of a set of frames from a video sequence (e.g., FIG. 3, steps 302, 304; spec., page 9, ¶ [0030]). For each frame in the selected set of frames, a set of motion vectors are determined and a motion angle for each motion vector is also determined (e.g., FIG. 3 steps 306, 308; spec., pages 9-10, ¶ [0031]). In each frame, at least two largest regions having substantially similar motion angles are identified (e.g., FIG. 3, step 310; spec., page 10, ¶ [0033]). The percentages of each frame covered by the at least two largest regions

and a statistical measure of the motion angles for at least one of the two largest regions are determined (e.g., FIG. 3, steps 312, 314; spec., page 11, ¶ [0035]). The percentages of each frame covered by the at least two largest regions and the statistical measure of the at least two largest regions are compared to threshold values to identify a pan or a zoom in the video sequence (e.g., FIG. 3, steps 318, 320; spec., pages 11-12, ¶ [0036]).

Grounds of Rejection to be Reviewed on Appeal

Whether claims 1-18 are unpatentable under 35 U.S.C. § 103(a) over PCT Application No. US97/08266 to Chang et al. ("Chang") and U.S. Patent No. 6,670,963 to Osberger ("Osberger").

Argument

<u>The Proposed Combination of Chang and Osberger Does Not Disclose "identifying at least two largest regions in each frame having substantially similar motion angles" and "determining percentages of each frame covered by the at least two largest regions" (claims 1-18)</u>

To establish a *prima facie* case of obviousness, the cited references, when combined, must teach or suggest all the claimed elements. *In re Royka*, 490 F.2d 981 (CCPA 1974); *see also In re Fine*, 837 F.2d 1071 (Fed. Cir. 1988) (reversing 35 U.S.C. § 103(a) rejection because examiner ignored material claimed limitation that was absent from reference). As the suggested combination of references fails to teach or suggest all of the limitations of the rejection claims, the Examiner's obviousness rejection was improper.

Specifically, the suggested combination of Chang and Osberger fails to disclose or suggest "identifying at least two largest regions in each frame having substantially similar

motion angles" and "determining percentages of each frame covered by the at least two largest regions" as variously recited in the independent claims.

Representative Claim 1 recites:

> A method of detecting at least one of a pan and a zoom in a video sequence, comprising:
>
> selecting a set of frames from a video sequence;
>
> determining a set of motion vectors for each frame in the set of frames;
>
> determining a motion angle for each motion vector;
>
> identifying at least two largest regions in each frame having motion vectors with substantially similar motion angles;
>
> determining percentages of each frame covered by the at least two largest regions;
>
> determining a statistical measure of the motion angles for at least one of the two largest regions; and
>
> comparing the percentages and statistical measure to threshold values to identify at least one of a pan and a zoom in the video sequence.

The claimed method detects the presence of a pan or a zoom in a video sequence. Initially, a set of frames are selected from the video sequence and a set of motion vectors are determined for each frame in the set of frames. For each motion vector, a motion angle describing motion vector orientation is then determined. At least two largest regions in each frame having motion vectors with substantially similar motion angles are identified and the percentage of each frame covered by the at least two largest regions is determined. To identified at least one of a pan and a zoom, a statistical measure of the motion angles for at least one of the identified two largest regions is computed and compared to threshold values.

Thus, the claimed method detects a pan or a zoom by identifying the two largest regions of each frame in a video sequence having substantially similar motion vector orientation. Motion angles are computed for each motion vector and the motion angles and used to identify

the largest regions having substantially similar motion vector orientation. By identifying the two largest regions of each frame having a substantially similar motion vector orientation, the claimed method allows detection of a pan or a zoom without computing global motion parameters (i.e., computing motion where most of the image points are uniformly displaced). Further, determining motion angles for each motion vector enables rapid identification of frame regions having substantially similar motion vector orientation by evaluating the motion angles. Determining a statistical measure for one of the largest frame regions, rather than the entire frame, reduces the computation necessary to detect a pan or a zoom in the frame, making pan or zoom detection more efficient.

In contrast, Chang merely discloses detecting moving objects within a frame by identifying areas of a frame having motion vectors different than other, non-moving, areas of the frame (Chang, page 17, lines 4-6). This detection merely compares motion vectors to a predetermined threshold value to eliminate areas of the frame with motion vectors below the predetermined threshold value (Chang, page 17, line 8). Chang further uses a linear transformation and a translation to identify moving and non-moving regions of a frame rather than to identify "at least two largest regions in each frame having motion vectors with substantially similar motion angles," as claimed. Further, as Chang does not identify "at least two largest regions in each frame having motion vectors with substantially similar motion angles," Chang cannot determine "percentages of each frame covered by the at least two largest regions," as recited in the independent claims. Because Chang detects all regions in a frame including motion vectors exceeding the predetermined threshold value, there is no determination of the "percentages of each frame covered by the at least two largest regions." Further, the

Examiner admits that Chang does not explicitly disclose determining percentages of each frame covered by the at least two largest regions. *See* Final Office Action dated May 2, 2008, page 4.

Osberger fails to remedy the deficient disclosure of Chang. Rather, Osberger discloses a segmentation algorithm that divides a video frame into a plurality of regions based on color and luminance (Osberger, Abstract). The segmentation algorithm processes both a current frame and a previous frame to produce motion vectors for the current frame (Osberger, 2:33-37). However, Osberger examines motion vectors associated with the complete current frame and the complete previous frame to generate an importance map for the current frame (Osberger, 3:23-30). There is no disclosure or suggestion in Osberger of "identifying at least two largest regions in each frame having substantially similar motion angles" or "determining percentages of each frame covered by the at least two largest regions," as claimed. Osberger merely examines a specified percentile of the camera motion compensated vector map to estimate motion in a scene. For example, Osberger examines the 98th percentile of the camera motion compensated vector map while disregarding the remaining 2 percent of the motion vectors to determine the amount of overall motion in a complete frame vector map (Osberger, 7:61-64). This m estimation does not identify "at least two largest regions in each frame having substantially similar motion angles" or determine "percentages of each frame covered by the at least two largest regions," but merely discounts a portion of the motion vectors in a complete frame during analysis. Hence, Osberger merely evaluates motion vectors from the frame as a whole, not motion vectors from different regions within the frame. The motion analysis disclosed in Osberger examines individual frames in their entirety and does not identify "at least two largest regions in each frame having substantially similar motion angles" or determine percentages of each frame covered by the at least two largest regions," as claimed.

While Osberger processes motion vectors for a current frame and for a previous frame to produce motion vectors for the current frame, the claimed invention identifies at least two largest regions within the same frame in a video sequence having substantially similar motion angles and computes the percentage of the frame covered by each of the identified largest regions. Further, Osberger takes "the mth percentile, such as the 98th percentile, of the camera motion compensated motion vector map"to estimate the amount of motion in a scene using a subset of the camera motion compensated motion vector map (Osberger, col. 7, lines 58-64). This motion vector map describes all motion in a complete frame, without identifiying different regions within the frame having syubstantially similar motion angles, and identifies areas in a scene where the viewer is likely to focus (Osberger, col. 8, lines 10-34). Nothing in Osberger indicates that the disclosed motion vector map or percintile analysis identifies "at least two largest regions in each frame having substantially similar motion angles" or determines the percentage of a frame covered by the two largest regions of the frame having motion vectors with substantially similar motion angles. Rather, Osberger identifies the total amount of object motion in a complete frame, regardless of where the motion occurs within the frame. Merely analyzing a specified percentile of a motion vector map does not determine of the percentages of a frame covered by the largest regions in the frame having motion vectors with substantially similar motion angles, but analyzes a subset of the motion vectors included in a complete frame, regardless of their location within the frame. Analyzing a specificed percentile of a compensated motion vector map is not equivalent to "determining percentages of each frame covered by the at least two largest regions," as the Examiner alleges. Taking a percentile of a motion vector map merely removes a certain percentage of the motion vector and does not identify regions having substantially similar motion angles.

In describing the rejection based on Chang and Osberger, the Examiner references Osberger to show the amount of motion activity in a histogram for frames having moving or non-moving backgrounds. *See* Final Office Action dated May 2, 2008, pages 2-3. The histogram referenced by the Examiner merely provides an indication of overall motion in a particular frame, but does not assocaite motion vectors with individual regions, such as the two largest regions, within each frame. However, Osberger does not disclose generating or analyzing a histogram, but discloses analyzing a specified percentile of motion vectors in a complete frame to determine the amount of motion in a scene. Contrary to the Examiner's assertions, Osberger makes no mention of a histogram, but merely discloses analyzing a specific percentile of motion within a frame. This percentile does not determine the percentage of each frame covered by the at least two largest regions, as there is no indication of which region of a frame includes the analyzed motion vectors.

The Examiner alludes to hypothetical cases of frames with moving backgrounds and frames with non-moving backgrounds to support application of Osberger to the claims. *See* Final Office Action dated May 2, 2008, page 2. Although a frame having a moving background would have different motion vectors than a frame having a non-moving background, Osberger does not disclose how analyzing a percentile of motion vectors in a complete frame identifies "at least two largest regions in each frame having motion vectors with substantially similar motion angles." Osberger merely analyzes an entire frame, and provides no disclosure of identifying regions within a frame having similar motion vectors. In the hypothetical examples provided by the Examiner, Osberger would be unable to identify particular regions within different frames having substantially similar motion angles, but would be able to broadly identify differences in total motion between frames having moving backgrounds and frames having non-moving

backgrounds. Determining the difference in overall motion between two complete frames does not determine "percentages of each frame covered by the at least two largest regions" having substantially similar motion vectors, as claimed.

Thus, the cited references, taken alone or in combination, do not disclose or teach the claimed invention. Therefore, claim 1 is patentable over the cited references and withdrawal of the rejection is respectfully requested.

Claims 7 and 13 similarly recite identifying "at least two largest regions in each frame having substantially similar motion angles" and determining "percentages of each frame covered by the at least two largest regions." Therefore, claims 7 and 13 are patentable over the cited references, both alone and in combination, for at least the same reasons discussed above with respect to claim 1.

In addition to reciting their own patentable features, claims 2-6, 8-12 and 14-18 variously depend from patentable base claims 1, 7 and 13. Accordingly each of dependent claims 2-6, 8-12 and 14-18 are also patentable.

The Examiner also rejected claim 16 as unpatentable over Chang and Osberger in view of Official Notice. However, the Official Notice relied upon by the Examiner does not overcome the deficiencies of Chang and Osberger. The Official Notice merely indicates that polar coordinates are a form of mathematical representation. However, this Official Notice does not disclose "determining percentages of each frame covered by the at least two largest regions" or "identifying at least two largest regions in each frame having motion vectors with substantially similar motion angles," as claimed. Therefore, the combination of Chang, Osberger and Official Notice fails to disclose the subject matter of claim 16.

Thus, claim 16 is patentably distinguishable over the cited references, both alone and in combination and withdrawal of the rejection is respectfully requested.

Conclusion

For the foregoing reasons, the Examiner's rejection of claims 1-18 was erroneous, and reversal of his decision is respectfully requested.

Respectfully submitted,
Adriana Dumitras and Barin G. Haskell

Dated: December 3, 2008 By: /Brian G. Brannon/

Brian G. Brannon, Reg. No. 57,219
FENWICK & WEST LLP
801 California Street
Mountain View, CA 94041
Tel: (650) 335-7610
Fax: (650) 938-5200
bbrannon@fenwick.com

Appendix: Claims Involved in Appeal

1. A method of detecting at least one of a pan and a zoom in a video sequence, comprising:

selecting a set of frames from a video sequence;

determining a set of motion vectors for each frame in the set of frames;

determining a motion angle for each motion vector;

identifying at least two largest regions in each frame having motion vectors with substantially similar motion angles;

determining percentages of each frame covered by the at least two largest regions;

determining a statistical measure of the motion angles for at least one of the two largest regions; and

comparing the percentages and statistical measure to threshold values to identify at least one of a pan and a zoom in the video sequence.

2. The method of claim 1, wherein the step of selecting a set of video frames from a video sequence further comprises:

identifying a scene cut between two frames in the video sequence; and responsive to the identification of a scene cut,

selecting a set of video frames from the video sequence that includes all the frames in the video sequence up to and including a frame just before the scene cut.

3. The method of claim 2, wherein frame differences and motion information are used to identify a scene cut.

4. The method of claim 1, wherein the motion angles are computed in one from the group of coordinate systems consisting of polar, Cartesian, spherical and cylindrical coordinate systems.

5. The method of claim 1, wherein the percentages of each frame covered by the at least two largest regions are determined from the number of pixels in each region as a percentage of the total number of pixels in a frame.

6. The method of claim 1, wherein the statistical measure is a variance.

7. A system for detecting at least one of a pan and a zoom in a video sequence, comprising:

a preprocessor for selecting a set of frames from a video sequence; and

a motion analyzer for determining a set of motion vectors for each frame in the set of frames, determining a motion angle for each motion vector;

identifying at least two largest regions in each frame having motion vectors with substantially similar motion angles determining percentages of each frame covered by the at least two largest regions, determining a statistical measure of the motion angles for at least one of the two largest regions, and comparing the percentages and statistical measure to threshold values to identify at least one of a pan and a zoom in the video sequence.

8. The system of claim 7, wherein the step of selecting a set of video frames from a video sequence further comprises

identifying a scene cut between two frames in the video sequence and responsive to the identification of a scene cut, and

selecting a set of video frames from the video sequence that includes all the frames in the video sequence up to and including a frame just before the scene cut.

9. The system of claim 8, wherein frame differences and motion information are used to identify a scene cut.

10. The system of claim 7, wherein the motion angles are computed in one from the group of coordinate systems consisting of polar, Cartesian, spherical and cylindrical coordinate systems.

11. The system of claim 7, wherein the percentages of each frame covered by the at least two largest regions are determined from the number of pixels in each region as a percentage of the total number of pixels in a frame.

12. The system of claim 7, wherein the statistical measure is a variance.

13. A computer-readable medium having stored thereon instructions which, when executed by a processor in a system for detecting at least one of a pan and a zoom in a video sequence, cause the processor to perform the operations of:

selecting a set of frames from a video sequence;

determining a set of motion vectors for each frame in the set of frames;

determining a motion angle for each motion vector;

identifying at least two largest regions in each frame having motion vectors with substantially similar motion angles;

determining percentages of each frame covered by the at least two largest regions;

determining a statistical measure of the motion angles for at least one of the two largest regions; and

comparing the percentages and statistical measure to threshold values to identify at least one of a pan or a zoom in the video sequence.

14. The computer-readable medium of claim 13, wherein the step of selecting a set of video frames from a video sequence further comprises:

identifying a scene cut between two frames in the video sequence; and responsive to the identification of a scene cut,

selecting a set of video frames from the video sequence that includes all the frames in the video sequence up to and including a frame just before the scene cut.

15. The computer-readable medium of claim 13, wherein frame differences and motion information are used to identify a scene cut.

16. The computer-readable medium of claim 13, wherein the motion angles are computed in polar coordinates.

17. The computer-readable medium of claim 13, wherein the percentages of each frame covered by the at least two largest regions are determined from the number of pixels in each region as a percentage of the total number of pixels in a frame.

18. The computer-readable medium of claim 13, wherein the statistical measure is a variance.

Evidence Appendix

None.

Related Proceedings Appendix

None.